Electronic Version
Stylesheet Version v1.1.1

Description

STRUCTURE AND METHOD FOR SILICIDED METAL GATE TRANSISTORS

BACKGROUND OF INVENTION

[0001] The invention relates to a semiconductor processing methods and structures, and more particularly to a structure and method for forming metal gates of transistors, especially for transistors in complementary metal oxide semiconductor (CMOS) technology.

[0002] Polysilicon gate electrodes have been a preferred material for the manufacture of gate electrode because of the special characteristics of polysilicon, particularly thermal robustness and the greater patterning precision available for etching a polysilicon gate. Many fabrication steps, such as annealing processes, require extremely high processing temperatures and therefore it is important to have a polysilicon gate withstands high temperatures during the fabrication process. Polysilicon gates are capable of withstanding high temperature processing of other elements

of transistors such as source and drain regions during dopant drive-ins. In addition, precise edges can be defined on polysilicon gates when etching a layer of polysilicon according to a resist pattern.

[0003] However, polysilicon gates have several disadvantages. Polysilicon is not a very good conductor of electricity, a quality which makes polysilicon transistors operate at slower speeds. Furthermore, a polysilicon gate is subject to the formation of a depletion region in operation in which charge carriers are depleted from the gate material above the gate dielectric. This varies from a metal electrode in which charge carriers remain plentiful throughout the electrode. The depletion region has the effect of making the gate dielectric appear thicker in operation than it actually is, such that more charge is needed to turn on the transistor having the polysilicon gate than the transistor having the metal gate. Consequently, in recent years, alternatives for replacing polysilicon gates have been considered.

[0004] Metals have long been a preferred material for manufacture of transistor gates since metals are better conductors of electricity, resulting in reduced gate contact resistance and faster device performance. However, manufacturing

of metal gates has been previously avoided because of difficulties in fabrication. For one, metal gates are not as thermally robust as polysilicon, responding poorly to high temperatures during processing of transistors or other elements of integrated circuits (ICs). In addition, metal gates cannot withstand the oxidation ambient present during some steps of gate fabrication such as gate sidewall spacer formation and/or gate sidewall oxidation. Furthermore, patterning accuracy required in gate formation is reduced when performing photolithography or other similar techniques on metal surfaces. Planar surfaces, which are a requirement for photolithographic patterning accuracy, are not easily obtainable in metals.

[0005] In recent years, there has been greater interest in the fabrication of metal gates, due at least in part to the development of a new approach. In this new approach, a sacrificial polysilicon gate is first formed, which is used during initial high temperature processing of the transistor and which is later replaced by metal gate structure. By this replacement gate process, initial severe transistor processing conditions such as high temperatures need not be modified, and the photolithography benefits associated with polysilicon processing are preserved. Moreover, the

initial use of a sacrificial polysilicon gate is benefited by the ability of polysilicon to block ion-implantation to the channel region of the transistor when performing source and drain implants to the transistor.

[0006] Metal gates are advantageously used in CMOS transistors including p-type field effect transistors (PFETs) and p-type field effect transistors (NFETs). In such transistors, the channel dopant profile requires a high dopant concentration near the gate dielectric with a sharp drop in concentration elsewhere in order for the gate to have good control over the channel. This requires that the transistors have a heavily doped shallow region under the channel. Such doping is typically obtained by implanting and annealing at a high temperature or by dopant drive-in at a high temperature.

[0007] A major challenge in fabricating such transistors is to perform the implant and anneal (or dopant drive-in at high temperature) in a suitable order relative to other steps used to form the transistors. The channel profile has to be established sometime after the sacrificial polysilicon gate is removed but before the metal gate is formed in its place. Adding to the complexity of transistor fabrication, a layer of silicide may be formed over the source and drain

regions of the transistor to enhance transistor performance. Silicides of some metals, particularly cobalt and nickel, are intolerant of high temperature processing, thus adding to the difficulties in fabricating such transistors.

[0008] These and other difficulties of fabricating metal gates present challenges to be overcome. Improved structures and fabrication methods are needed that can address current fabrication challenges and enhance the overall performance of metal gates.

Summary of Invention

[0009] According to an aspect of the invention, a structure and method are provided for fabricating a field effect transistor (FET) having a metal gate structure. A metal gate structure is formed in an opening within a dielectric region formerly occupied by a sacrificial gate. The metal gate structure includes a first layer contacting a gate dielectric formed over a semiconductor region of a substrate. The first layer includes a material selected from the group consisting of metals and metal compounds. The gate further includes a silicide formed over the first layer. The FET further includes a source region and a drain region formed on opposite sides of the metal gate structure. Preferably, the source region and drain region are sili-

cided. The silicide on the source and drain regions is preferably formed by the same process in which the silicide is formed over the first layer of the metal gate structure.

[0010] According to a preferred aspect of the invention, the silicide layer is formed in a self-aligned manner from a layer of silicon overlying the first layer of the metal gate structure and a layer of metal deposited over the layer of silicon. An underlying layer of silicide may be formed at an interface between the first layer and the layer of silicon.

Brief Description of Drawings

[0011] Figures 1 though 13 are cross-sectional views illustrating stages in a processing method according to an embodiment of the present invention.

Detailed Description

[0012] The present invention provides a way of forming a silicide on the source and drain regions of PFETs and NFETs by a self-aligned technique after forming metal gates of such transistors which does not damage the structure of the metal gate.

[0013] The present invention particularly addresses and solves problems associated with the fabrication of metal gates

according to the prior art. In particular, the present invention provides transistors having a metal gate wherein a self-aligned silicide is formed on the source and drain regions and over a first layer of the metal gate structure. A feature of the invention is that the salicide is simultaneously formed over the metal gate, thus avoiding process complexity but without adversely affecting the characteristics of the metal gate.

[0014] A feature of the invention is the performance of the silicidation step at a time after processing at high temperatures (i.e. 500 deg. C and above) such that the silicide layer is not damaged by high temperature processing. The silicidation step is conducted in a manner to avoid damage to the gate during subsequent etching procedures. These and other features of the invention will now be discussed in greater detail as follows.

[0015] Figure 1 is a sectional view illustrating a stage in a fabrication method according to an embodiment of the present invention. Figure 1 illustrates a semiconductor substrate 100. The term "substrate" is used herein for ease of reference and includes a variety of types of substrates including bulk semiconductor substrates, semiconductor-on-insulator substrates such as "silicon-on-insulator"

(SOI) substrates, germanium (Ge) substrates, strained silicon/silicon germanium (SSi/SiGe) substrates and silicon germanium (SiGe) substrates. Such substrates include a region of a single-crystal semiconductor at a main surface thereof. The term substrate may also be applied to a substrate having a thin deposited semiconductor layer when the process described herein is used to form a thin film transistor (TFT).

[0016] Isolation structures as shown at 110 are then formed on the substrate 100. These isolation structures 110 can include a variety of structures such as shallow trench isolation and are formed selectively on the substrate 100. The area of the substrate defined by and in between the isolation structures 110 is known as the active area and is shown in Figure 1 at 120. The active area 120 will house the active electrical devices. The purpose of the isolation structures is to provide electrical isolation between devices in various and particularly adjacent active areas 120.

[0017] Figure 2 illustrates a next processing stage in which initial process steps are conducted to form a sacrificial polysilicon gate on the substrate 100. As shown in Figure 2, an etch stop layer 200 is formed by deposition or grown on the substrate 100. In a preferred embodiment, the etch

stop layer 200 includes an oxide such as silicon dioxide. Alternatively, the etch stop layer may include a nitride such as silicon nitride, silicon oxynitride or other similar material. As shown in Figure 2, a layer 210 of polysilicon is then deposited as a sacrificial gate material on the etch stop layer 200.

[0018] As shown in Figure 3, the sacrificial gate material 210 and the etch stop layer 200 are patterned together. Since the sacrificial gate is formed of polysilicon, as is widely used in NFET and PFET transistors, patterning processes are readily available.

[0019] Photolithographic imaging followed by etching is used to pattern the gate stack 300 including the etch stop layer (illustratively provided as an oxide layer 200) and the polysilicon 210. Etching can be performed by an anisotropic etch process such as a reactive-ion plasma etch (RIE), for example.

[0020] In the next processing step, spacers 400 are formed on the sidewalls of the gate stack structure, as shown in Figure 4. In a preferred embodiment, the spacers 400 are formed of a nitride such as silicon nitride. Preferably, L-shaped spacers 410 are provided underlying the spacers 400. The L-shaped spacers are preferably formed by de-

positing a first layer of material, preferably a layer of oxide, on gate stack 300 and substrate 100. A second layer of material is then deposited over the first layer of material. The second layer of material is preferably a nitride such as silicon nitride. An anisotropic vertical etch such as RIE is then used to remove the first and second layers of material except where they collectively coat sidewalls of the sacrificial gate, to produce the structure including spacers 400 and 410 that is shown in Figure 4.

[0021] Figure 4 also illustrates a process of forming the source and drain regions as shown at 420. Raised source and drain regions are now formed by selectively growing an epitaxial layer of silicon in areas 420 on the surface of the substrate 100. During this step, the sacrificial gate 210 and spacers 400, 410 prevent the epitaxial layer from being grown onto the channel region 430 of the substrate 100.

[0022] Thereafter, ion implantation into the source and drain regions is performed separately to define NFETs and PFETs. During this step, areas of the substrate 100 in which PFETs are to be formed are masked while source and drain regions of the NFETs are implanted. Likewise, areas of the substrate in which NFETs are formed are masked when the

source and drain regions of the PFETs are implanted.

[0023] The regions shown at 440 in Figure 4 can also be implanted to form lightly doped source/drain extension regions and/or halo implanted regions. The extensions as shown at 440 and halo implantation can be formed by dopant ion implantation before spacer 400 is formed. In addition, either or both n-type and p-type impurities can be implanted as needed to form the specific desired components. In a preferred embodiment where complementary metal oxide semiconductor (CMOS) technology is used, n-type dopants and p-type dopants must be implanted into respective portions of the substrate to form the source and drain regions of the NFETs and the PFETs. A sufficiently thick layer of polysilicon has the ability to block ion implantation into underlying layers. The sacrificial polysilicon gate 300 and the spacers 400, 410 function together as an implant mask during ion implantations to form the source and drain regions.

[0024] The use of a sacrificial polysilicon gate in these early stages of processing a metal gate allows high temperature processing to be performed. For example, a high temperature dopant drive-in process is generally required following implanting a dopant into source and drain regions

of the substrate after patterning the gate.

[0025] Next, as shown in Figure 5, an interlevel dielectric layer 500 is blanket deposited over the substrate 100. In a preferred embodiment, the interlevel dielectric material includes silicon dioxide (SiO_2). Preferably, the interlevel dielectric 500 is planarized, stopping at the sacrificial polysilicon 210. A variety of processes are available for planarization such as chemical mechanical polishing (CMP).

[0026] As shown in Figure 6, the sacrificial gate 210 is then removed, as by RIE, stopping on the etch stop layer 200 so as to avoid damaging the surface of the substrate 200 under which the conduction channel 630 of the transistor will be formed. As a result, as shown in Figure 6, an opening 600 is formed, bound by the etch stop layer 200 and the sidewalls 610 of the L-shaped spacers 410.

[0027] Once the sacrificial gate 210 has been removed, a shallow threshold voltage adjustment implant can be performed through the etch stop layer 200. Thereafter, an anneal is performed to diffuse the dopant ions to the desired distribution and to repair damage to the crystal structure of the substrate 200 as a result of the implant. Then, the etch stop layer 200 is removed, as by a dry directional etch se-

lective to silicon and to nitride such as RIE.

[0028] Alternatively, a doped glass drive-in process is used to provide the threshold voltage adjustment to the transistor's conduction channel. In such process, the etch stop layer 200 is removed, as by a dry directional etch such as RIE. A thin layer of dopant material such as arsenic doped glass (if an n-type implant is desired) or borosilicate glass (if a p-type implant is desired) is then deposited onto the surface of substrate 100 in the place of the etch stop layer 200 within the opening 600, as shown at 700 in Figure 7. Thereafter, the dopant is distributed to the desired depth and distribution through a drive-in anneal. The doped glass material is then removed from the opening 600, as by a dry directional etch.

[0029] Whichever process is used to provide the threshold voltage adjustment, the surface of the substrate at 700 is now cleaned and the final gate dielectric 800 is now formed in the opening 600, as shown in Figure 8. In an embodiment, the gate dielectric 800 includes an oxide layer and is thermally grown on the substrate 100 within the opening 600. In another embodiment, the gate dielectric 800 is formed by deposition, as by low pressure chemical vapor deposition (LPCVD) of a material such as silicon dioxide, silicon

nitride, or silicon oxynitride. Other choices of materials exist for the gate dielectric. For example, a gate dielectric of hafnium oxide (HfO_2) or of zirconium oxide (ZrO_2) can be formed as a gate dielectric having a desirably high dielectric constant K, higher than that of either silicon dioxide, silicon nitride or silicon oxynitride. Such high-k gate dielectric may be advantageous for a particular application, such as where a thicker gate dielectric is needed to protect against dielectric breakdown but without sacrificing transistor switching performance.

[0030] Figure 9 illustrates a stage in formation of the metal gate structure including a metal gate layer 900. In the embodiment illustrated in Figure 9, tungsten (W) is used as a preferred material for a metal gate layer 900. One or more metals and/or compounds of metals can be used to form the metal gate layer 900 such that the metal gate layer 900 may include only one or more metals, or only one or more compounds of metals, or, alternatively a combination of metal(s) and metal compound(s). In an example, the metal gate layer 900 may preferably include a metal such as tungsten which has a workfunction at about the middle bandgap of silicon. In another example, the metal gate layer 900 may include tungsten silicide. While the

layer 900 may include a compound of a metal, reference will be made hereinafter to a "metal gate layer." The metal gate layer 900 is first deposited in the opening shown previously as 600. Then, in a preferred embodiment, the metal gate layer is planarized to the level of the interlevel dielectric 500 using chemical mechanical polishing (CMP).

[0031] Next, processing to form silicides in the source and drain regions and contacting the gate is performed. The present invention provides a way of forming self-aligned silicides in the source and drain regions of the transistors at a timing that does not interfere with other required processes or the quality of the completed transistors. As discussed above, in forming p-type field effect transistors (PFETs) and n-type field effect transistors (NFETs), it is desirable to perform shallow implants into the channel regions of the transistors in order to adjust the threshold voltages to desirable levels. High temperature processing is required in order to anneal the channel regions of PFET and NFET transistors following the shallow threshold voltage adjustment implants. However, if the source and drain regions of the transistors are silicided prior to such high temperature processing, device degradation can result. Many annealing processes require temperatures of over 800 de-

grees Celsius. Many silicides, including nickel silicide (NiSi), are unstable above temperatures of 500 degrees Celsius. Therefore, such silicides should not be formed prior to completing such high-temperature processing.

[0032] Consequently, it is desirable to form a silicide in the source and drain regions of the transistors after implanting and annealing the channel region. However, an opportunity to silicide the source and drain regions after annealing the channel region is not presented until after the metal gate is fully formed. This presents a problem. If the silicide is to be formed by a self-aligned technique, a deposited metal has to reacted with the silicon in contact therewith to form the silicide. The excess unreacted metal must then be subsequently removed, i.e. etched away selective to the silicide and dielectric materials that are present. However, unless the metal of the underlying metal gate is somehow protected, this etch to remove the unreacted metal could damage the metal gate.

[0033] These concerns are addressed by the present invention as follows. An intervening layer of silicon is formed over the metal gate layer 900 as a way of protecting the metal gate layer 900 such that the silicide can be formed thereafter using a self-aligned technique. Thus, in the next stage of

processing, illustrated in Figure 10, the metal gate layer 900 is recessed within opening 600 to provide space between spacers 410 for the formation of a silicon layer 1000. In a preferred embodiment, the metal gate layer 900 is recessed by removing, preferably 20 nm to 50 nm of metal. The metal gate layer 900 is preferably recessed using dilute hydrogen peroxide (H_2O_2) but other techniques may be utilized as well, such as a dry etch technique.

[0034] After recessing the metal gate layer 900, a doped silicon layer 1000 is then deposited over the metal gate layer 900. In a preferred embodiment of the present invention, the layer of silicon 1000 comprises a layer of polysilicon (p-Si) or amorphous silicon (a-Si). In one such embodiment the p-Si or a-Si is then doped in-situ, using either n-type or p-type dopants, to reduce the gate resistance but without affecting the threshold voltage (Vt) of the transistor including the gate. The silicon layer 1000 is then subsequently lowered to the level of the interlevel dielectric 500, preferably through planarization by chemical mechanical polishing (CMP).

[0035] The process by which the silicide is formed is now described, with reference to Figures 11 through 13. Figure

11 illustrates a further stage in processing during which the interlevel dielectric 500 is removed to expose the source/drain regions 420. As illustrated in Figure 12, a second metal layer 1200 is conformally deposited. The second metal layer 1200 is deposited in a manner so that a thin layer of metal is deposited on the substrate 100, the spacers 400 and the doped silicon layer 1000. The second metal layer 1200 preferably includes cobalt or nickel. Alternatively, a metal such as platinum or titanium could be used instead. An anneal is then performed to react the metal layer 1200 with the silicon of the source and drain regions 420 and the silicon layer 1000 of the gate to form a self-aligned silicide 1300. In such embodiments, when the deposited second metal layer 1200 is cobalt, the annealing process to form $CoSi_2$ as the silicide 1300 is usually conducted within a temperature range of 600 to 750 degree Celsius. Alternatively, when the deposited second metal layer 1200 is nickel. The annealing process to form NiSi as the silicide 1300 is conducted within a temperature range of 400 to 500 degrees Celsius. When the deposited second metal layer is titanium, the annealing process to form $TiSi_2$ as the silicide 1300 is conducted within a temperature range of 500 - 750 C. When the de-

posited second metal layer is platinum, the annealing process to form PtSi as the silicide 1300 is conducted within a temperature range of 450 – 800C.

[0036] When the metal gate layer 900 is a metal which reacts with silicon to form a silicide, the annealing process also forms a silicide layer between the metal gate layer 900 and the doped silicon layer 1000. In embodiments in which tungsten is used as the metal gate layer 900, this results in the formation of a tungsten silicide (WSi_2) layer.

[0037] Subsequent to the annealing process, a selective etch is performed to remove excess unreacted metal from all areas except where a layer of silicide is formed. In such manner, excess metal 1200 is removed from areas contacting the spacers 400. The resulting structure is shown in Figure 13.

[0038] As shown in Figure 13, a metal silicide 1300 is provided on the source and drain regions 420 on the substrate 100. There is also a layer of silicide 1300 on top of the doped silicon 1000. A silicide layer 1310 is also formed between the metal gate 900 and the doped silicon layer 1000. The doped silicon layer 1000 will be hereinafter referenced as intervening layer.

[0039] During annealing to form the self-aligned silicide 1300,

the doped silicon layer 1000 may actually be fully consumed by reactions with a metal of the underlying metal gate layer 900 and with the second metal layer 1200 to form the two silicide layers 1310 and 1300. In such case, the two silicide layers 1310 and 1300 merge with each other.

[0040] In still another embodiment, when the metal gate layer 900 is a silicide material or a material that does not form a silicide in contact with silicon, the silicide layer 1310 will not appear after annealing. In such embodiment, depending on the relative quantities of material in the doped silicon layer 1000 and the second metal layer 1200, the annealing process to form the silicide 1300 may or may not fully consume the doped silicon layer 1000.

[0041] While the invention has been described in accordance with certain preferred embodiments thereof, those skilled in the art will understand the many modifications and enhancements which can be made thereto without departing from the true scope and spirit of the invention, which is limited only by the claims appended below.